



The Effect of Data Structure and Model Choices on MFA Results: A Comparison of Phosphorus Balances for Denmark and Austria

Klinglmair, Manfred; Zoboli, Ottavia; Laner, David; Rechberger, Helmut; Astrup, Thomas Fruergaard; Scheutz, Charlotte

Published in:
Resources, Conservation and Recycling

Link to article, DOI:
[10.1016/j.resconrec.2016.03.009](https://doi.org/10.1016/j.resconrec.2016.03.009)

Publication date:
2016

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Klinglmair, M., Zoboli, O., Laner, D., Rechberger, H., Astrup, T. F., & Scheutz, C. (2016). The Effect of Data Structure and Model Choices on MFA Results: A Comparison of Phosphorus Balances for Denmark and Austria. *Resources, Conservation and Recycling*, 109, 166-175. <https://doi.org/10.1016/j.resconrec.2016.03.009>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

1

2

3

4

5

6 **The effect of data structure and model choices on**

7 **MFA results: a comparison of phosphorus balances for**

8 **Denmark and Austria**

9

10

11

12 Manfred Klinglmair¹⁾, Ottavia Zoboli²⁾, David Laner²⁾, Helmut Rechberger²⁾, Thomas Fruergaard

13 Astrup¹⁾, Charlotte Scheutz¹⁾

14

15

16 1) Technical University of Denmark, Department of Environmental Engineering

17 2) Vienna University of Technology, Institute for Water Quality, Resource and Waste Management

18

19

20

21 **KEYWORDS:** substance flow analysis; MFA; phosphorus; data quality; uncertainty assessment; data

22 reconciliation

1 **ABSTRACT**

2 Material Flow Analysis (MFA) studies for a particular substance often exist for several different
3 countries or regions, but share a similar goal and scope. In direct comparisons of such regional resource
4 budgets, the importance of the choices made in establishing an MFA system tends to be disregarded.
5 We identify and quantify the effects of choices made in system layout, data material and uncertainty
6 assessment on the outcome of regional MFAs using two recent country-scale MFAs (of Denmark and
7 Austria) of phosphorus as a case study.

8 We highlight the differences in system boundaries and definition of flows and processes. We quantify
9 types and choice of data sources; analyse the consistency of the data used by looking at the extent of
10 data reconciliation, as a measure of model quality; quantify the effect of different approaches to
11 uncertainty assessment; and show the influence of aggregating/disaggregating flows.

12 We show that differences in system layout are mostly attributable to varying goals and scope
13 definitions. Direct comparison of uncertainties across studies is problematic: both studies draw on
14 similar types of data sources, yet they show very different uncertainty assessments; the uncertainty
15 assessment in MFA is always subjective to a certain extent. We demonstrate that reconciliation of
16 conflicting data provides a useful measure to assess data consistency and model quality: data are more
17 consistent (5% average change in reconciled data) in the Austrian than in the Danish (9%) case. We
18 suggest an iterative approach to uncertainty assessment. Likewise, we demonstrate the effect of the
19 aggregation of flows on model uncertainty. These findings quantify and emphasise the importance of
20 examining MFA studies' metadata and suggest an approach to be followed when drawing on such
21 studies as a source of information.

22

1. Introduction

Material Flow Analysis (MFA; Baccini & Brunner 2012) has become a widespread approach to visualise the anthropogenic turnover of materials across a defined geographical area. For a number of substances, e.g. phosphorus (P), several MFA studies—often called Substance Flow Analysis (SFA) when studying a substance—of regional or national households with similar goal and scope definitions are available. Patterns of resource use are analysed to evaluate resource efficiency and recycling activities (e.g. Chen and Graedel 2012), anthropogenic material stocks are investigated using dynamic analysis to quantify current and future resource potentials (e.g. Müller et al. 2014), and sources and sinks of problematic substances are the focus of MFA studies addressing environmental pollution issues (e.g. Hansen and Lassen 2003). While MFA studies may obviously differ with respect to their geographical scope, the authors' modelling choices and the database used for MFA are a different set of factors often disregarded in the comparison of regional resource budgets. Country or multi-country level MFA studies are quite consistent with respect to system boundaries, and external trade and national accounts, across substances and countries (Fischer-Kowalski et al. 2011). For P, Seyhan (2009) suggests a standard template regarding processes, flows and system boundaries, for P resource budgets on the national or regional level.

However, there is considerable freedom in data selection and model choices. This implies that the model and data structure of an MFA typically vary from one study to another, an issue tending to get little mention in comparisons of P flow studies, which are often taken at face value (e.g. the review of phosphorus MFAs by Chowdhury et al. 2013). Data uncertainty quantification and consequent uncertainty analysis has become an integral part of MFA, and has recently been analysed in more detail by Laner et al. (2014 and 2015b). Schwab et al. (2015) have moreover presented a detailed framework to systematically characterise data used in an MFA. Still, uncertainties due to model structure and

scenario assumptions (or arbitrary choices of the practitioner), and their influence on MFA results, are rarely addressed and therefore difficult to analyse.

It is the aim of this work to examine the effect of the MFA system layout and modellers' choices on the outcome of regional MFAs, using the country-scale MFAs of P in Denmark and Austria as case studies. The Danish study (Klinglmair et al. 2015) examines the anthropogenic P budgets of three distinct regions in Denmark as subsystems of the P flows on the national scale. The Austrian study (Zoboli et al. 2015a) presents a 22-year timeline of detailed P budgets on the national level with particularly detailed accounts of flows in the waste management sector. Both studies share the objective of examining anthropogenic P flows from a standpoint of resource efficiency for countries of comparable size and population.

The following characteristics of an MFA model are of interest to the above objective: 1) system boundaries, layout (i.e. processes/flows shown, resolution, definition of processes, flow definitions), 2) types and choice of data sources, 3) the closely linked issues of data reconciliation and uncertainty assessment: over-determination, i.e. the difference between the number of independent balance equations and unknown variables (flows) of the model; the results of different approaches to quantify uncertainties in both studies; the extent of data reconciliation as a measure of data consistency and model quality; and the effects of aggregation/disaggregation of flows on model uncertainty. Points 1 and 2 are discussed in qualitative terms; point 3 consists of systematically quantifiable results. It is not the aim of this study to compare the P budgets of the two countries as such; instead, we take the specific examples of these two studies to illustrate, on a more general level, the extent of differences that can occur in the areas above due to the MFA practitioner, the data available, and the model formulation, apart from actual physical differences between the systems. In this way, we demonstrate a series of steps to analyse MFA studies in order to make plain, and quantify, this often-overlooked

distinction. While this work is based on studies of P, the conclusions we draw may be applied to a wider range of country- or regional-scale resource budgets.

2. Materials & methods

2.1 Source material and background

The present study is based on two recent P balances (Klinglmair et al. 2015; Zoboli et al. 2015a) for Denmark and Austria, conducted at the Technical University of Denmark and Vienna University of Technology, respectively. Table 1 shows characteristics pertinent to the two countries' P budgets. The area used for agriculture is similar in both countries, but relatively smaller in Austria, due to the larger area of the country and the high share of forests. However, more livestock is produced in Denmark, with a particularly high number of pigs compared to Austria.

1 *Table 1* Basic country data pertaining to P households in Austria (AT) and Denmark (DK).

		AT	DK
Population		8,200,000	5,580,000
Area	ha	8,387,899	4,309,400
Agricultural area	ha	2,879,000	2,567,000
	% of total	34%	60%
Forest area	ha	3,405,752	600,427
	% of total	41%	14%
Livestock	Cattle	1,976,527	1,567,970
	Pigs	3,004,907	12,931,678
	Sheep	361,183	143,889
	Horses	81,637	61,476
	Goats	72,358	(no data)
	Poultry	14,644,413	18,206,943
Livestock/ha agricultural land	Cattle	0.7	0.6
	Pigs	1.0	5.0
	Sheep	0.13	0.06
	Horses	0.02	0.02
	Goats	0.03	(no data)
	Poultry	28.8	7.1

2

3

4

5

6

7

8

9

10

11

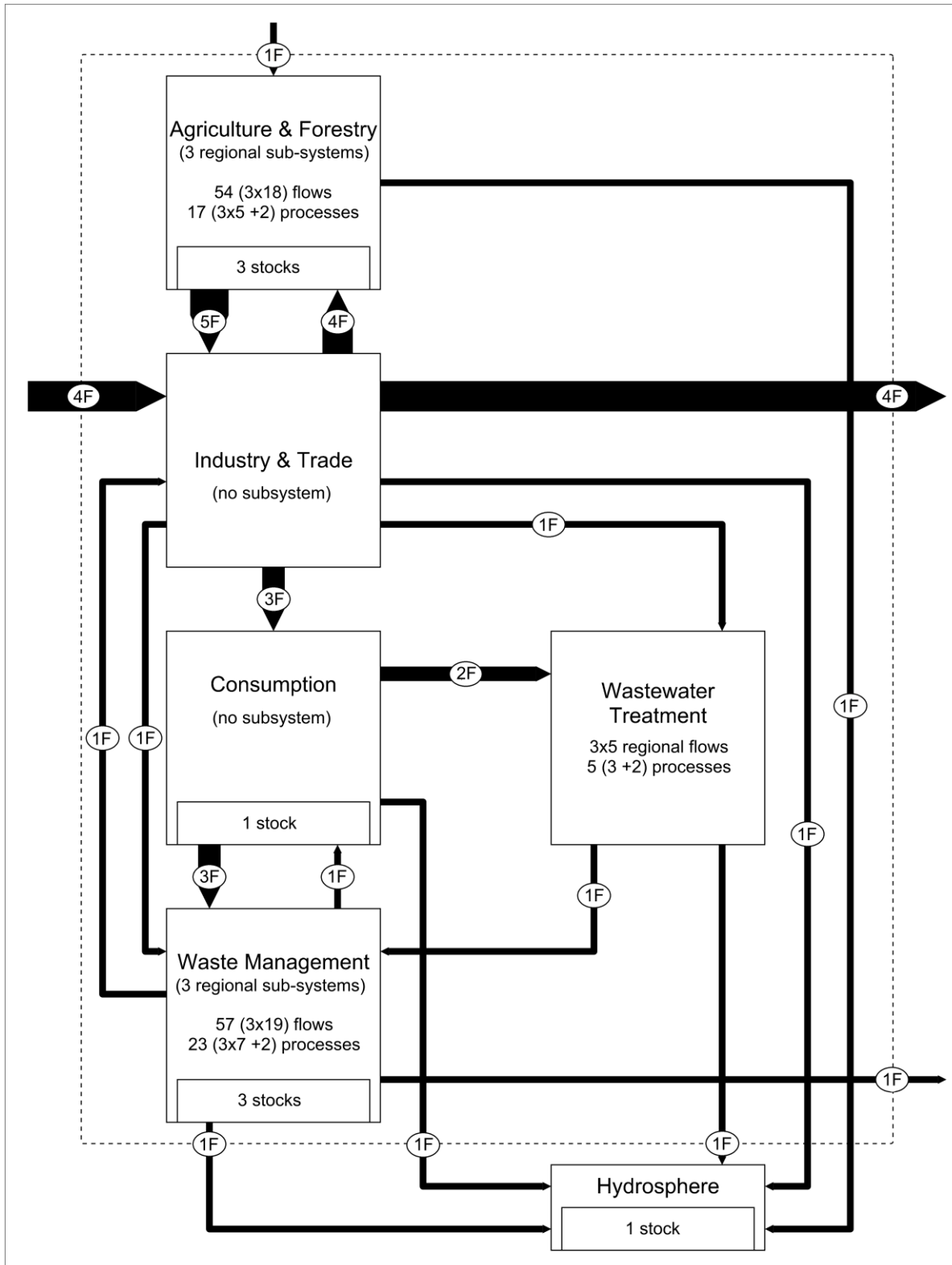
12

13

While the geographical scope and the areas under scrutiny are comparable, the objectives of the studies differ in their respective focus. In the Austrian case, Zoboli et al. (2015a) draw upon an existing P budget (Egle et al. 2014) to examine the utility of multi-year timelines in MFA studies, highlighting considerable variation in the period from 1990–2011. The model layout was only slightly modified from the earlier study, and comprises 56 processes and 122 flows, including 10 ‘virtual’ processes to aggregate or split flows for visualisation purposes. 11 processes contain subsystems, where the process is further disaggregated into flows and processes on a sub-system level. The Danish study (Klinglmair et al. 2015), conducted for the year 2011, aims for a higher spatial resolution by incorporating three regional subdivisions, for the agriculture and waste management processes, into the national P balance. The purpose of this study was to investigate imbalances in the P household and potential for P recovery

1 resulting from differences in agricultural practice, and population and industrial density between the
2 Danish regions. The model layout comprises a total of 50 processes and 166 flows, including 12
3 'virtual' processes and 9 processes containing subsystems (see Tables S.1 and S.2 in the supplementary
4 data for flows and data sources). For the purpose of this comparison, we look at the results for the year
5 2011 in both studies. Both systems are overdetermined, i.e. more variables (flows) are known than are
6 necessary to solve the balance equations. In the Danish model, all 166 flows are determined, while all 7
7 stock changes are unknown and therefore calculated using mass balance equations. This means that
8 there exist 7 unknowns for 41 independent balance equations (processes not containing a sub-system).
9 In the Austrian case, 16 of 122 flows and 7 out of 8 stock changes are calculated, so that we find 44
10 independent balance equations and 23 unknowns. The degree of overdetermination is the number of
11 independent balance equations (constraints) minus the number of unknown variables; the Danish model
12 can therefore be said to be overdetermined to a higher degree. Figure 1 shows simplified qualitative
13 representations of both MFA systems. It has to be noted that several processes shown in the main
14 system in the Austrian study (*animal husbandry, crop farming, forestry & miscellaneous soils*, and
15 *bioenergy*; Figure 1b) are found in sub-systems in the *agriculture* process in the Danish study (see
16 Figure 1a).

17



1

2 *Figure 1a*

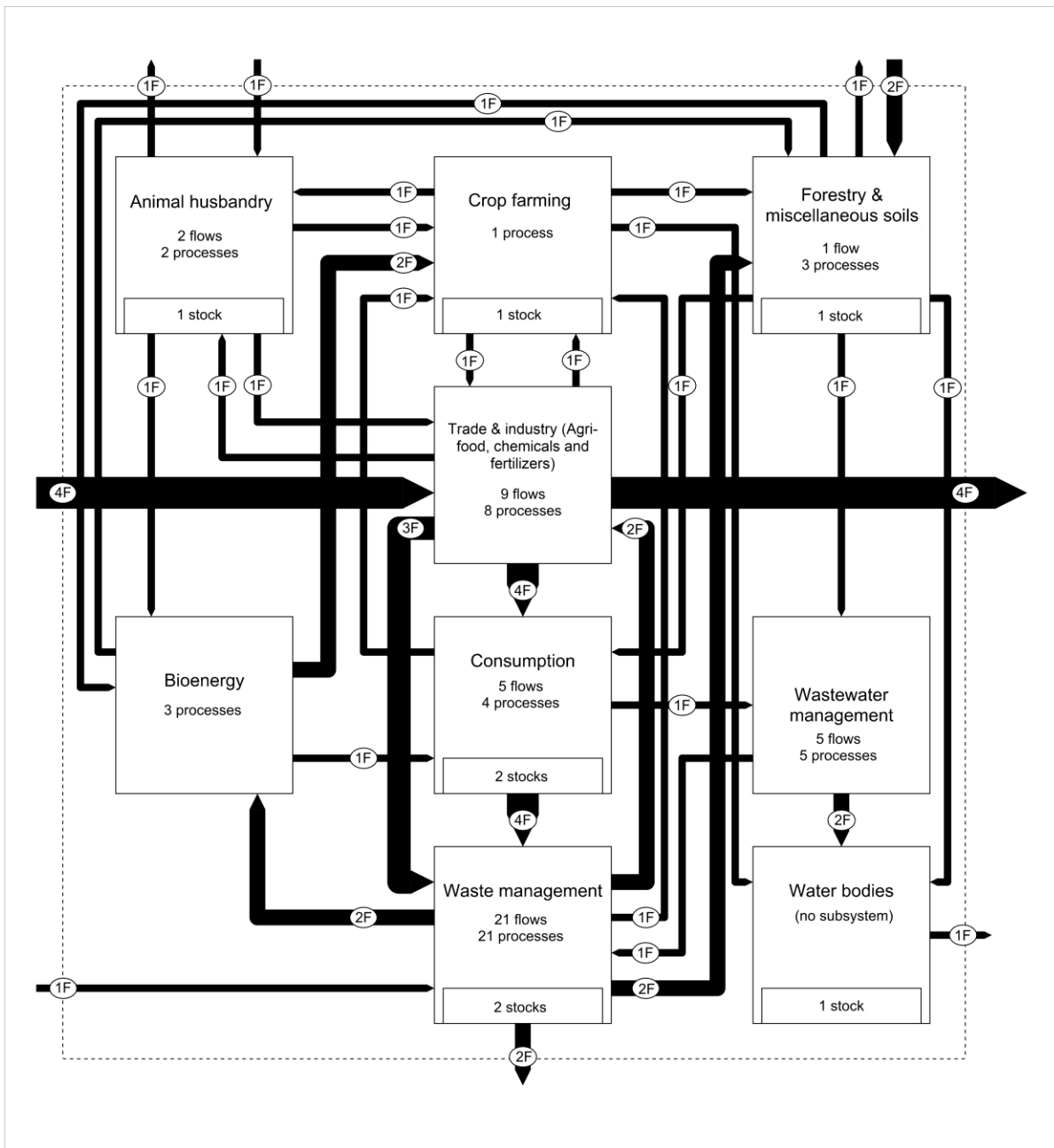


Figure 1b

Figure 1. Simplified qualitative representation of the MFA systems for P flows and stocks in Denmark (a) and Austria (b). Main systems are shown, i.e. without subsystems contained in the processes. The number of processes, flows, and stocks in the subsystems are listed below the process name. The

1 numbers of flows and processes in each of the 3 regional subsystems are shown separately in Figure 1a
2 (Denmark), where applicable. Arrows are marked (e.g. 1F, 4F ...) with the number of flows they
3 represent, with arrow width proportional to the number of flows. The dashed lines denote the system
4 boundary. Processes pertaining to crop farming, forestry, animal husbandry, and biogas are found in
5 subsystems in the Danish study.

6 7 8 2.2 Material Flow Analysis (MFA)

9 Both studies use the STAN 2.5 software (Cencic & Rechberger 2008) for balancing material
10 flow systems with options of data reconciliation and uncertainty propagation, assuming all uncertain
11 flows to be random, normally distributed variables given by mean value and standard deviation. Figure
12 1 shows a simplified graphical representation of the two MFA models.

13 Since flow values are based on data of varying quality (and hence uncertainty) from multiple
14 sources, these data will generally not be fully consistent with the model, but contradict each other to a
15 certain degree, i.e., the input and output flows of processes will not balance completely. This problem
16 can be solved by data reconciliation as implemented in STAN. Data reconciliation in STAN is a
17 weighted least-squares optimisation of the measurement adjustments subject to model constraints
18 (system balances, where the weights are chosen to be the inverse of the measurement variance; see
19 Laner et al. 2014). A further result of data reconciliation is the reduction of the uncertainty of the
20 reconciled flows.

21 22 2.3 Approach

As a *first* step, attention needs to be paid to the differences in the definitions of system layout and system boundaries, even as conventions about boundaries at various geographical scales have been developed (see Chowdhury et al. 2013). Nevertheless, the overall system layout warrants attention in any comparison of MFA studies. Furthermore, a (mostly qualitative) evaluation of the essential differences in the system layouts in representing the respective P flows and stocks, and in the definitions of processes and flows, is necessary.

Second, we examined the data material on which the models were built. We categorised the data sources used in the quantification of P flows and stocks (material/bulk flows as well as P concentrations) for both systems and ranked them according to their relative importance. Since flow values entered in the model are generally based on several sources, we looked at how frequently data sources from each respective category were cited as references for flow values.

Third, we compared the consistency of the models, their layouts and the approaches to uncertainty assessment. ‘Consistency’ here means the fit of the data given the mass balance constraints of the model in consideration of the defined uncertainty ranges for the input data. The more the entered values contradict each other, the higher is the conflict in the database in view of the mass balances of the model. The effect of data reconciliation can be quantified as the change of individual flow values after data reconciliation in STAN (in % of the entered values); the mean of all deviations gives a measure of the extent of data reconciliation or ‘fit’ of the model and data (see Eq. 1; see also Zoboli et al. 2015a), with \bar{D} being the mean relative deviation of the reconciled data from the entered values, m the total number of flows, x_n the entered value for flow n , and \hat{x}_n the reconciled value for flow n .

$$\bar{D} = \frac{\sum_{n=1}^m \left| \frac{x_n - \hat{x}_n}{x_n} \right| * 100}{m} \quad (1)$$

1 *Fourth*, the quantification of uncertainty in MFA studies, even if reproducible and systematic, is
2 subjective to a certain degree (see Laner et al. 2015b), which puts limitations on a direct comparison of
3 uncertainty ranges alone. We therefore compare and evaluate the approaches to uncertainty assessment
4 as well as the quantitative effects thereof, in conjunction with data reconciliation in STAN.

5 *Finally*, we evaluate the effects of aggregation/disaggregation of flows: often, a flow is an
6 aggregate of several smaller flows (as is often the case in the Danish study, for example, where certain
7 flows on the national scale are disaggregated into regional flows). This aggregate flow is either known
8 and it may be subject to data reconciliation, such as other determined flows; or, the flow is unknown
9 and is calculated as the sum of the flows it is composed of. In the former case, data reconciliation will
10 result in a reduction of uncertainties, because more information is given (i.e. the constituent flows are
11 known as well as the total (=aggregated) flow, hence, resulting in over-determination). In the latter
12 case, the uncertainties of the constituent flows are propagated to derive the uncertainty of the total
13 (aggregated) flow. Depending on how many aggregate flows are present in a model and whether they
14 contribute to overdetermination or not, this may have a considerable effect on the degree of data
15 reconciliation and the consequent uncertainty of the model results.

16 17 **3. Results & discussion**

18 19 **3.1 Qualitative aspects**

20 21 *3.1.1 System layout*

22 The system boundaries for both studies could be expected to be defined in similar ways. Indeed,
23 they differ only with regard to water bodies. The Danish study places the hydrosphere outside the

1 system boundary, considering P flows to water bodies simply as losses. The Austrian study, on the
2 other hand, includes water bodies in the system; this can be attributed to the slightly different goal &
3 scope definitions of the two studies. Whereas the Austrian study put a focus on P as a major contributor
4 to eutrophication in addition to the P resource management issues (this is also highlighted by a follow-
5 up study on the effect of anthropogenic P emissions on P load dynamics in the river Danube; Zoboli et
6 al. 2015b), the Danish study put emphasis mainly on the resource aspects of P management in
7 Denmark. Figure 2 shows the summed in- and outflows, and stock changes, of the main processes in
8 the Danish and Austrian models. Note that ‘Agriculture & forestry’ has been combined in the figure,
9 although in the Austrian case it is divided into three processes in the main system. The high turnover in
10 the industry & trade process results from the process handling most imports and exports, as well as
11 most exchanges with the consumption and agriculture and forestry processes.

12

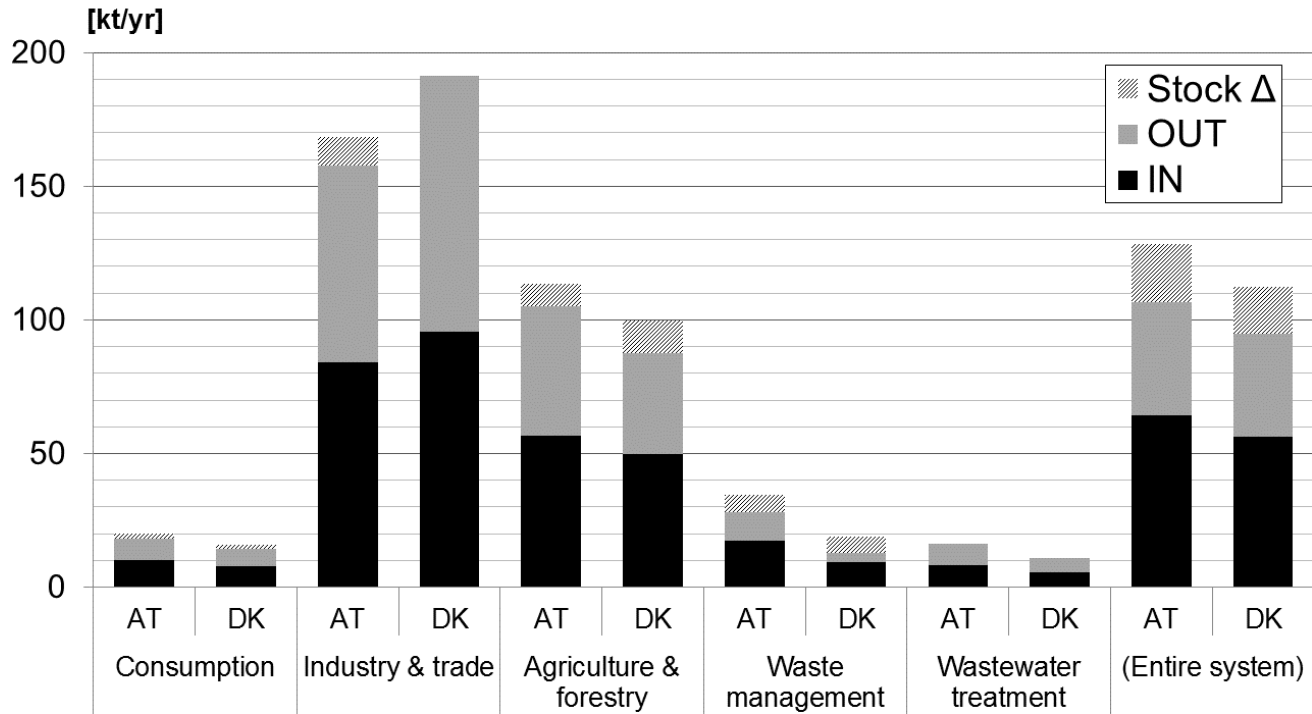


Figure 2. Overview of summed in- and outflows and stock changes per main process, for both studies examined, shown in kt/yr (AT: Austria; DK: Denmark).

When compared to the template for country-scale P balances put forward by Seyhan (2009), the main systems of both studies conform to the suggested types of processes, flows and stocks. Exceptions are mineral P reserves (non-existent in both countries) as well as a *Bioenergy* process in the main system for Austria (not part of the template; of similar magnitude in Denmark, but located in the *Agriculture* subsystem), and minor differences in the classification of various wastewater flows. *Wastewater treatment* is the only main process without a stock in both systems. Minor differences between the two studies can be more frequently found in the detailed subsystems, and process definitions. *Wood & paper industry* is found in the *Forestry* process in the Austrian case; in the Danish case, any industry is located in the *Industry and trade* process, with wood from forestry (which plays a considerably smaller role in Denmark) subsumed under the *Agriculture* process; see Figure 1. A second

1 such example is animal processing: slaughtering is not identified as a separate process in either study,
2 thus the activity is not explicitly located in the model.

3 The availability or use of specific datasets is linked to the layout of the MFA model. This is
4 highlighted by the use of datasets provided by national statistical offices regarding agricultural
5 production (with one of the main turnovers of P). In the Austrian study, the datasets used were national
6 supply balances; the Danish study draws on production data for domestic agricultural production. The
7 statistics referenced by the Austrian study list national data only; the corresponding Danish statistics
8 give regionalised data. Other, regionalised datasets also exist for Austria (e.g. Statistics Austria 2015).
9 These differences in model layout were therefore not caused by data availability. However, the datasets
10 used would not readily allow the reproduction of a regionally differentiated model layout, such as the
11 Danish one. Table 2 compares the datasets by national statistical offices from which the respective
12 values were drawn. In cases where regionalised statistical data for agricultural production exist in the
13 Danish case, they are reflected in the model. Moreover, these datasets do not list private production of
14 vegetables in the kitchen gardens for Denmark. The flow ‘garden vegetables’ exemplifies the influence
15 of data availability on model layouts: the flow is not shown in the Danish model (i.e. it is located
16 within the *Consumption* process), whereas in the Austrian case, a subsystem of consumption shows
17 production from private gardens (see Figures 1 a & b; see also Table S.1). A direct comparison of the
18 MFA models’ layouts can yield insight into such data gaps, or otherwise draw attention to erroneously
19 omitted (‘forgotten’) flows.

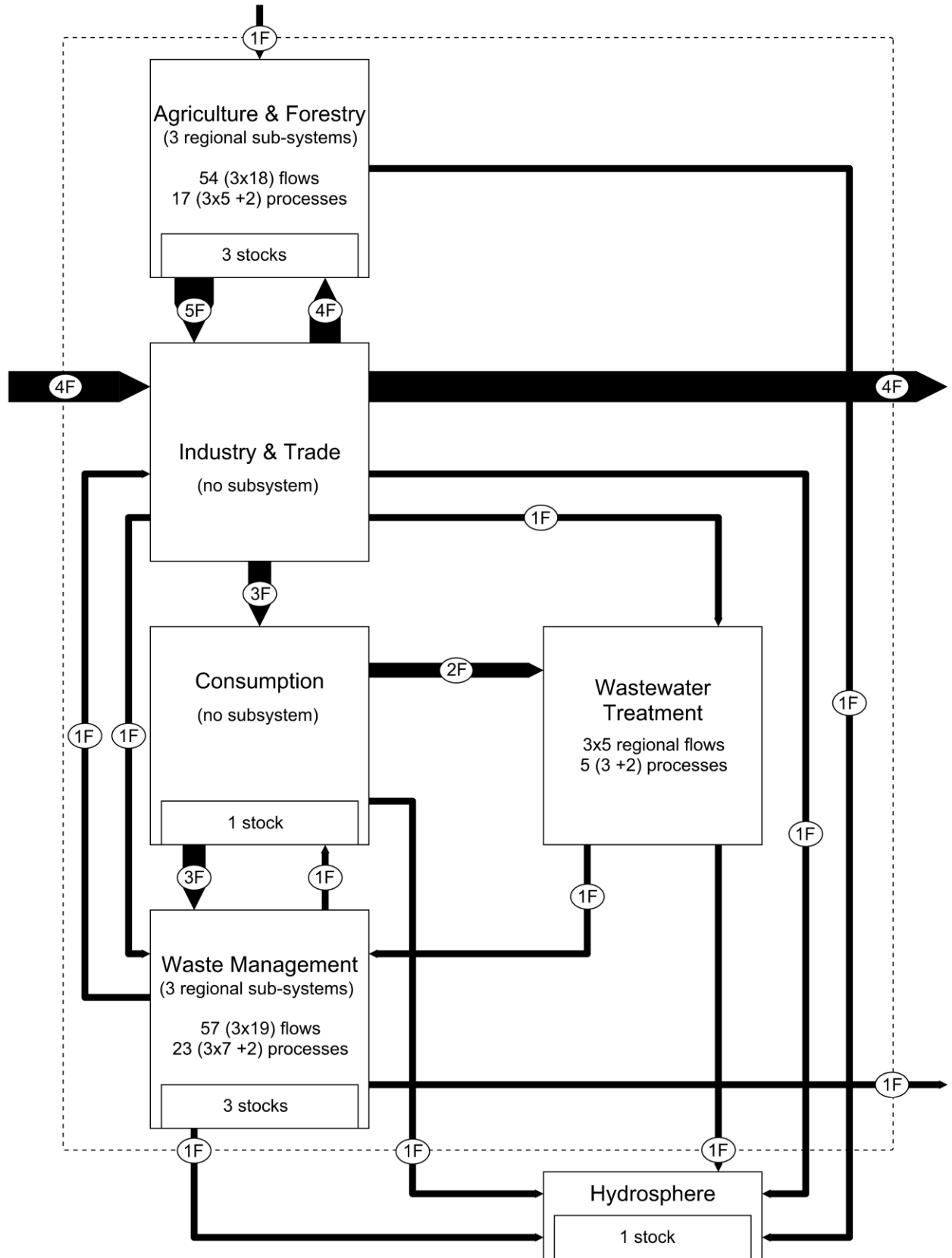


Table 2. Comparable flows based on datasets from national statistical offices (see Tables A.1 and A.2 for references). The flow ‘Garden vegetables’ has no counterpart in the Danish study.

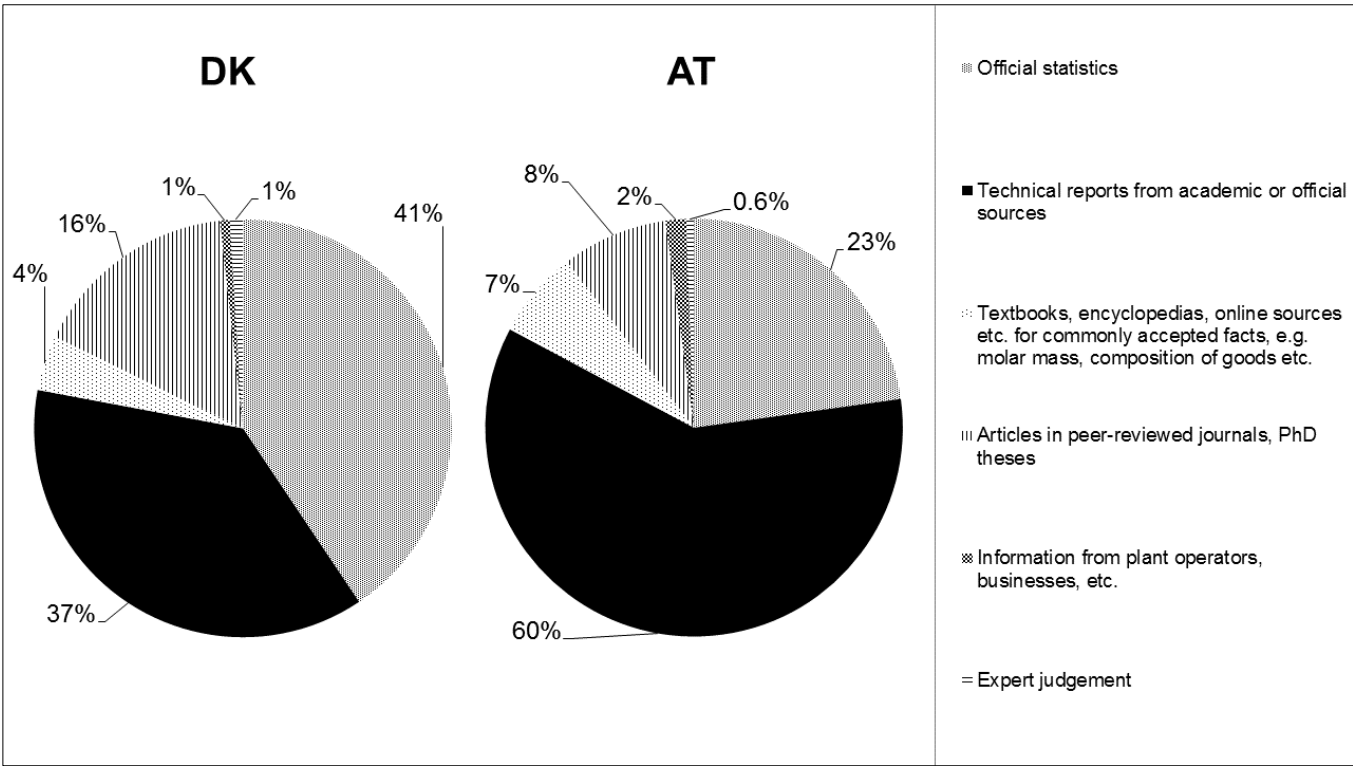
AT			DK		
Flow name	Regionalised data?	Regionalised flows?	Flow name	Regionalised data?	Regionalised flows?
Agricultural products	no	no	Plant products (food)	yes	yes
Seeds	no	no	Seeds & planting material	yes	yes
Food	no	no	Food	no	no
Import food	no	no	Imported food	no	no
Export food	no	no	Exported food	no	no
Eggs and milk production	no	no	Milk & eggs	yes	yes
Meat production	no	no	Meat & slaughterings	yes	yes
Garden vegetables	no	no	-	-	-
Import live animals	no	no	Other imports	no	no
Export live animals	no	no	Other exports	no	no

Without a clear justification in the respective studies, and if data availability is similar, such differences in subsystems can be reasonably attributed to the authors’ judgement. When compared to each other, and held against earlier templates such as those shown by Seyhan (2009) and Chowdhury et al. (2013), there is evident consensus on the basic system boundaries and layouts in P MFAs.

A notable aspect, however, is the higher spatial resolution of the model layout through the multi-regional aspect in the Danish case, leading to a multiplication (by the number of regions considered) of all flows quantified on the regional level. This engenders higher requirements on data consistency for these flows. In the example at hand, the resolution in terms of different *types* of flows (i.e. of different compositions/functions) is lower in the Danish than in the Austrian case, but regional sub-division results in a higher overall number of flows (and the resulting higher risk of data conflicts between connected flows).

3.1.2 Data sources

1 We divided the data sources used in the MFA models into 6 categories (Figure 3). Typically,
 2 flow values entered into STAN are based on data from more than one source. Figure 3 gives an
 3 overview of these sources, showing the relative frequency of references from each category being cited
 4 as sources for flow values; this includes both references cited for mass (bulk) flows of materials and P
 5 concentrations. In total, 399 citations are given in the Danish and 500 in the Austrian study. Generally,
 6 flow values are based on several data sources; for Figure 3, each single reference for each flow was
 7 counted. Both studies rely strongly on (usually non-peer-reviewed) technical reports from either
 8 academic or official sources as well as data from statistical offices, together comprising approximately
 9 80% of the data sources in both cases. The Danish study draws more strongly upon official statistics.
 10



11
 12 *Figure 3. Relative use of data sources in both systems (by the total number of citations). For a*
 13 *complete list of flows with corresponding references, see supplementary data (Tables S.1 and S.2).*

3.2 Quantitative discussion

3.2.1 Data consistency

Figure 4 shows an evaluation of data consistency per process: the smaller the deviation of a flow value after data reconciliation from the entered value, the better the consistency of the input data given the balance constraints of the model. Thus, the consistency can provide an indication of the ‘quality’ of the mass balancing. A grouping around processes makes sense in order to identify possible ‘weak spots’ in the system. In the Danish model, the average deviation of flows, as a result of data reconciliation over all flows, amounts to 8.6%, ranging from a 4.4% average of flows connected to, or contained in subsystems within, the *Consumption* process to 8.9% for *flows connected to the Industry and trade* process. For 25% of the flows, the extent of data reconciliation is stronger than average. In the Austrian model, the average deviation across all flows is 4.9%, with flows pertaining to *Wastewater treatment* being the lowest at 2.2%, and flows connected to and contained in *Industry and trade* being the highest at 7.7%, with 30% of flows showing stronger than average data reconciliation. In light of this, the Austrian study shows a slightly higher overall consistency in mass balancing. Additionally, a higher degree of overdetermination may result in a higher extent of data reconciliation, since it creates more possibilities for data conflicts; the Danish model is overdetermined to a higher degree than the Austrian model, particularly because of the aggregation of regional flows into known total flows on the national level (see Section 2.1).

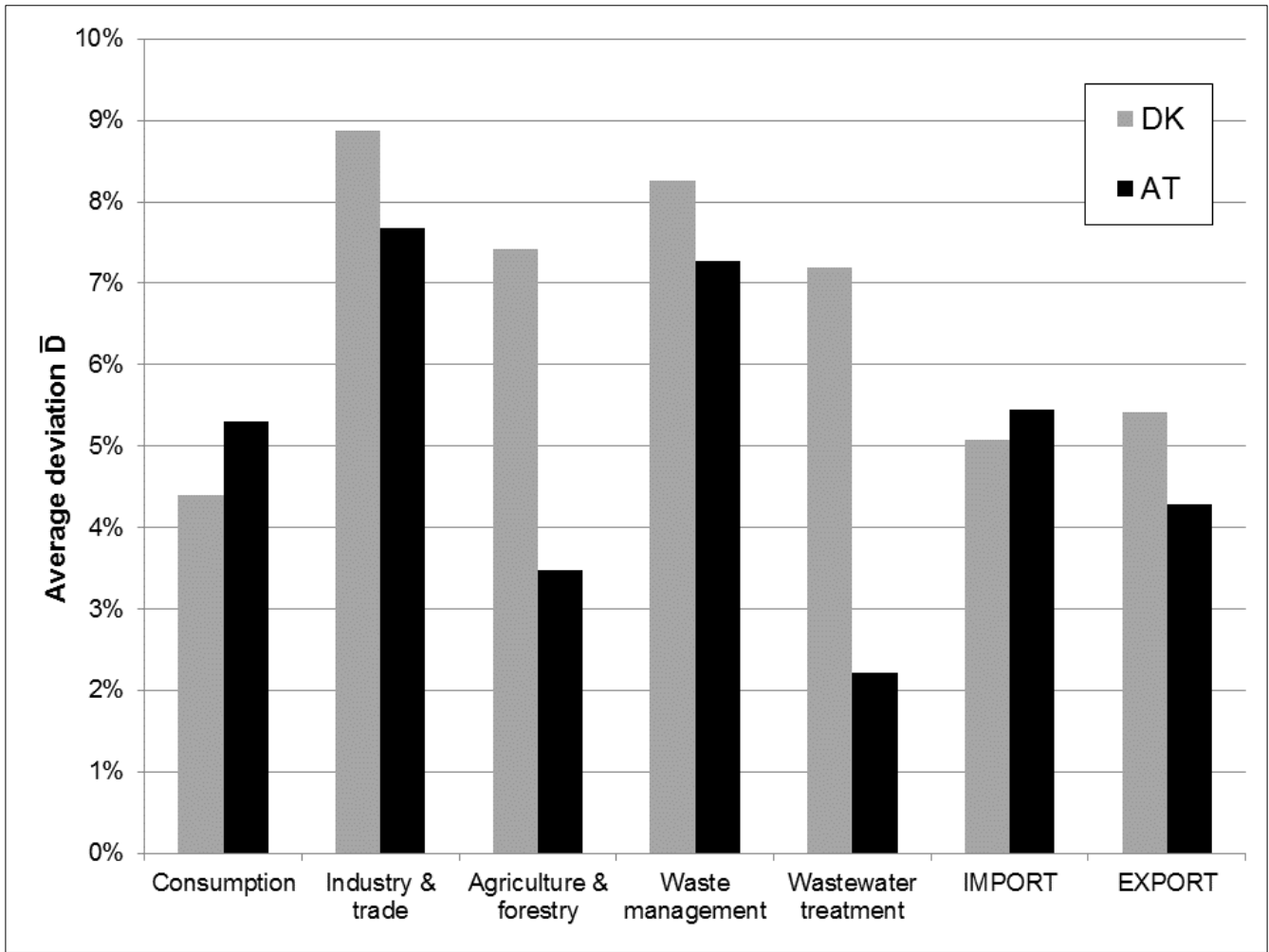


Figure 4. Deviation of reconciled from entered flow values per process in the main system and associated subsystems. Each column shows the average deviation \bar{D} (see Section 2.3, Eq. 1), in % of the entered values, of all flows connected to or belonging to a subsystem in a process. In the Austrian case, 'Agriculture & forestry' consists of 3 discrete processes (Animal husbandry, Crop farming, and Forestry & miscellaneous soils) in the main system that are grouped together here; for Denmark, such differentiation takes place in subsystems.

The extent of measurement adjustments in data reconciliation is influenced (weighted) by the uncertainties of entered data (see Section 2.2). Hence, the more heterogeneous the uncertainties of

individual flows throughout the model, the more reconciliation will concentrate on the most uncertain flows. It has to be noted that this limits, to some extent, the explanatory power of an average value of data reconciliation to measure consistency. The more heterogeneous the uncertainties of flows (the larger the difference between the smallest and highest), the higher will be the changes in the few very uncertain flows, while less uncertain flows change very little. On average, these changes may be similar to a system where uncertainties are more homogeneous, and may obscure these underlying differences.

3.2.2 Uncertainty assessment

The Danish study draws upon the uncertainty concept developed for MFA by Hedbrant & Sörme (2001). Depending on the data source, all the data are assigned a specific uncertainty level. Klinglmair et al. (2015) use 5 such levels; each uncertainty level corresponds to a category of data sources (e.g. data from national statistical offices, information from plant operators, etc.). Each uncertainty level is associated with an uncertainty factor, which is used to calculate asymmetric uncertainty ranges for the input data. Since STAN relies on the input data given as the mean value and standard deviation (i.e. symmetric ranges), the Danish study draws upon the adaptation of this approach by Laner et al. (2015b), converting these factors into coefficients of variation (CV). In this adaptation, the product of the mean value and the uncertainty factor is defined as the mean value plus two standard deviations. Thus, the symmetric interval around the mean corresponds to a 95% confidence interval. Based on this approach, an uncertainty level (and corresponding CV) is assigned to both material (bulk) flows and P concentrations therein, resulting in one CV for the P flow. Calculating the uncertainty of a P flow shown in the model from two uncertain values in this way resulted in 11 different CVs (uncertainty ranges) for characterising the uncertainty of flow values in the model.

1 The Austrian study makes use of a more refined approach for data quality assessment and
2 uncertainty characterisation (Laner et al. 2015a) based on the pedigree matrix by Weidema & Wesnæs
3 (1996), developed to express uncertainty due to data imperfection in Life Cycle Assessment (LCA).
4 The approach quantifies data uncertainty via a matrix assigning discrete scores, expressed as
5 coefficients of variance, to each of the data quality indicators: *reliability*, *completeness*, *composition*,
6 *temporal correlation*, *geographical correlation*, and *further correlation*. CVs are determined based on
7 the evaluation scores, and the overall CV of each datum is calculated as the square root of the summed
8 squares of the individual CVs. Due to the combination of various indicator scores for each datum, a
9 greater variety of uncertainty ranges is obtained for the P flows of the Austrian model in comparison to
10 the Danish model.

11 While both methods to quantify uncertainties follow a well-defined procedure, in their choice of
12 uncertainty levels associated with the respective quality of data sources, they are to some degree based
13 on the modellers' best judgement and assessment of data quality. This makes it difficult to compare the
14 uncertainty levels of corresponding flows directly between the two systems: a type of flow may occur
15 in both systems and be based on the same data source, yet have a different uncertainty range assigned
16 in each model. Both studies show a similar reliance regarding which types of data sources they draw
17 upon (see Section 3.1.2). However, the distribution of the uncertainty ranges assigned to flows
18 throughout the model differs considerably. The relative majority of flows are assigned relative standard
19 deviations of 30–50% in the Danish study; for Austria, the relative majority of flows are assigned
20 relative standard deviations of 10–20% (Figure 5). The different distributions of uncertainty ranges in
21 Figure 5 clearly indicate that the uncertainty assessment is not directly comparable between the studies,
22 regardless of a largely similar composition of the data material (Section 3.1.2).

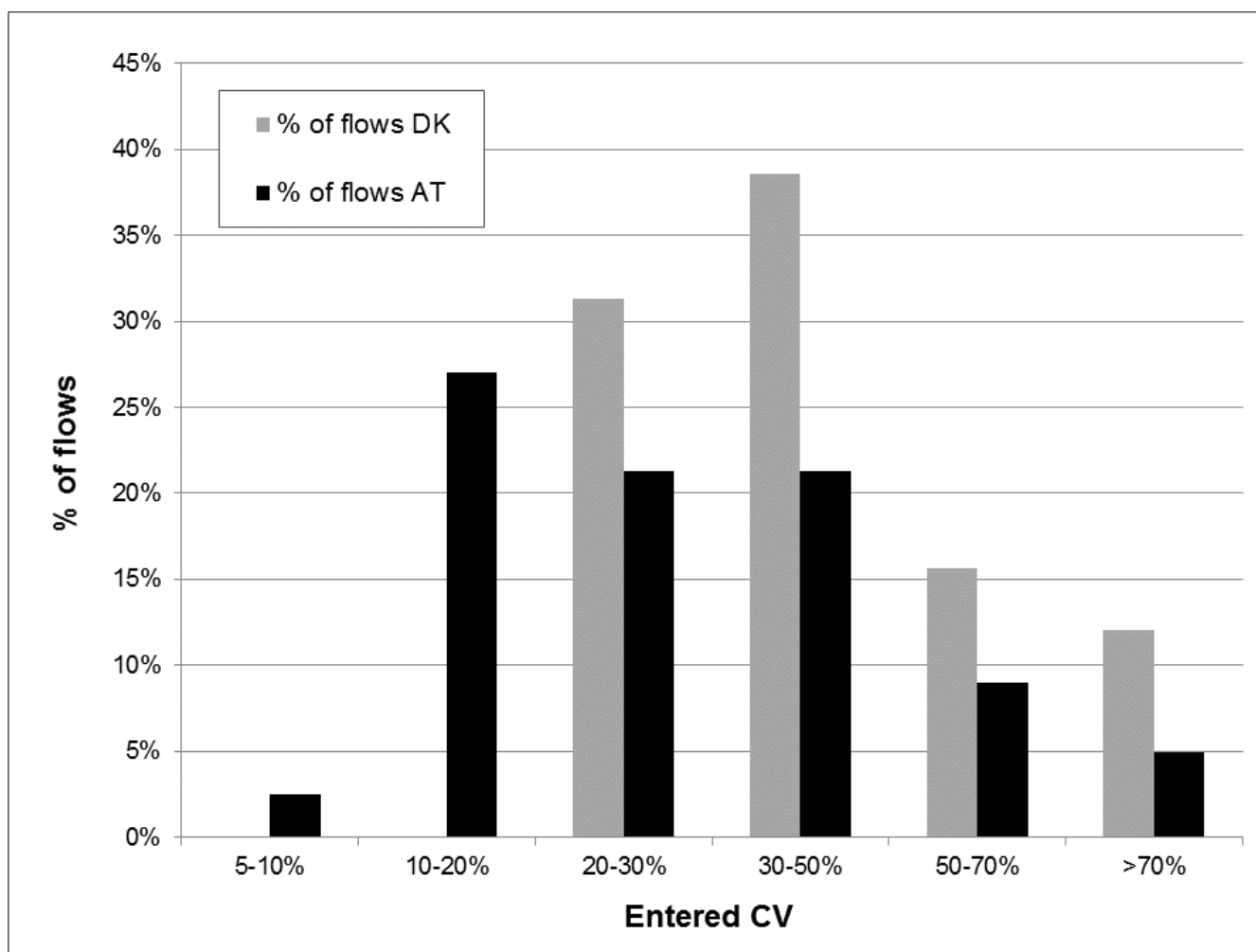


Figure 5. Distribution of uncertainty ranges (coefficient of variation, or CV) as entered in the model: percentage of flows with CV of 5–10%, 10–20%, 20–30%, 30–50%, 50–70%, >70%, respectively.

A high confidence in the quality of input data, and assigning low uncertainty ranges (see Figure 5) to a majority of flows as a consequence, will have a detrimental effect on the consistency of the mass balancing if there is substantial conflict between the inserted mean values for the flows. In such cases, overly strong reconciliation (i.e. the reconciled value is associated with a very low probability, given the probability density function of the input flow) may occur and the data quality assessment, the quantitative uncertainty estimates, and/or the model equations would need to be critically evaluated.

Moreover, data reconciliation in STAN leads to a reduction of uncertainty throughout the system (see Section 2.2); Figure 6a shows the effects on the two systems we examined. In both cases, relative uncertainties after data reconciliation are generally lower, with the exception of 38 flows in the Danish and 34 flows in the Austrian model. The most notable exceptions occur for highly uncertain flows, where reconciliation resulted in much higher relative uncertainties for individual flows. (Data reconciliation always results in a reduction of absolute uncertainty; an increase in relative terms can only be due to a decrease in the flow value).

The further to either side of the normal distribution—as defined by the uncertainty range entered by the user—a reconciled value lies, the less probable it becomes given the original information. Figure 6b compares the entered flow values with the reconciled values. If data reconciliation results in changes of flow values larger than the defined uncertainty range, it is a very strong indication that the original uncertainty characterisation was overly optimistic, and that the reconciliation is too strong. Conversely, if the uncertainty ranges defined by the modeller are much larger than the changes of flow values resulting from data reconciliation, such a pessimistic determination of data quality results in a loss of model precision. Seen in conjunction with Figure 6b, this shows that the uncertainty assessment in both studies resulted in similar uncertainty ranges; the Austrian study, however, is somewhat more optimistic regarding the flows with the smallest uncertainties. In order to ensure a quantification of uncertainties corresponding to the extent of data reconciliation, and to avoid overly optimistic (and pessimistic) uncertainty estimates, an iterative approach to uncertainty (cf. Laner et al. 2014) is advisable. The main goal of such a process is to ensure that all necessary changes of flow values remain within the specified uncertainty ranges.

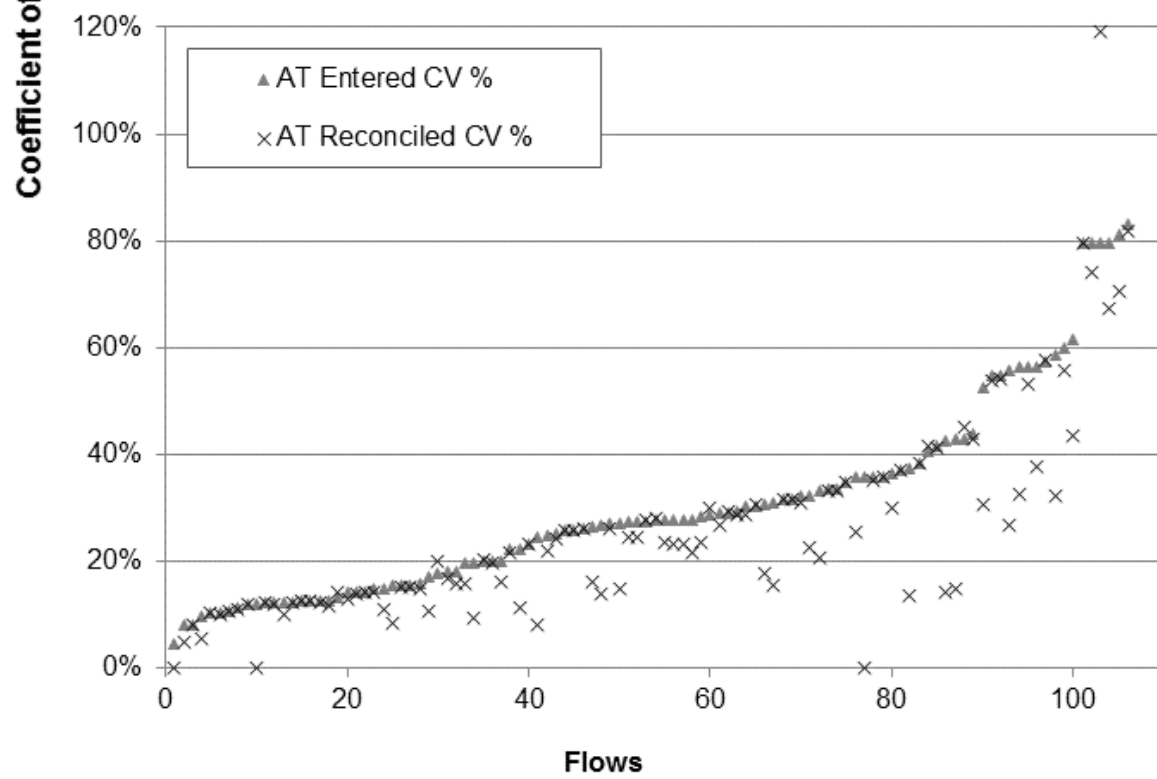
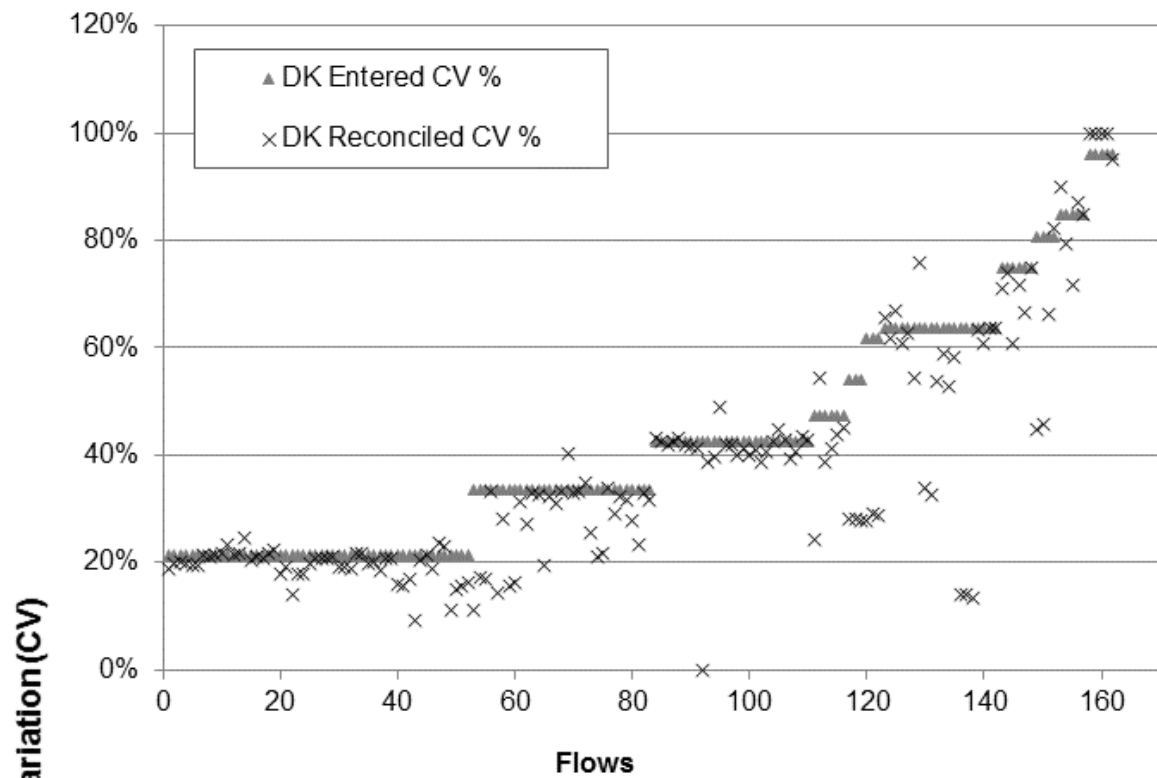


Figure 6a

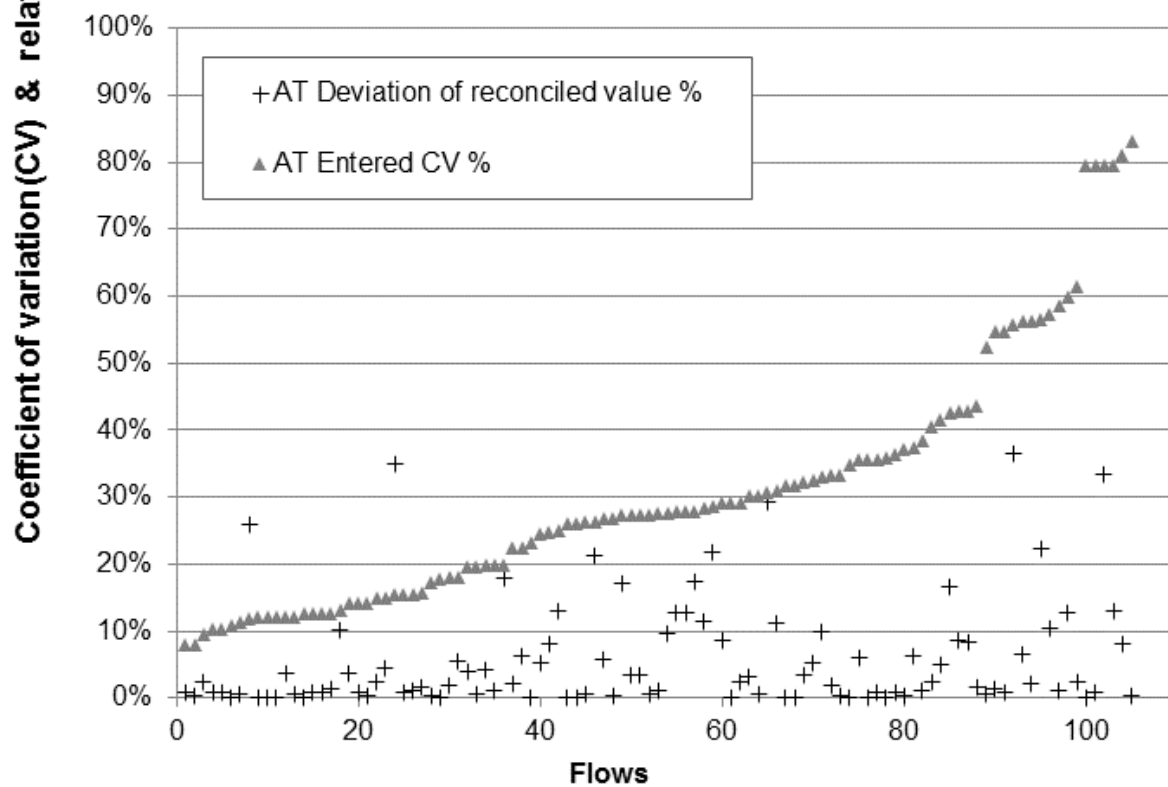
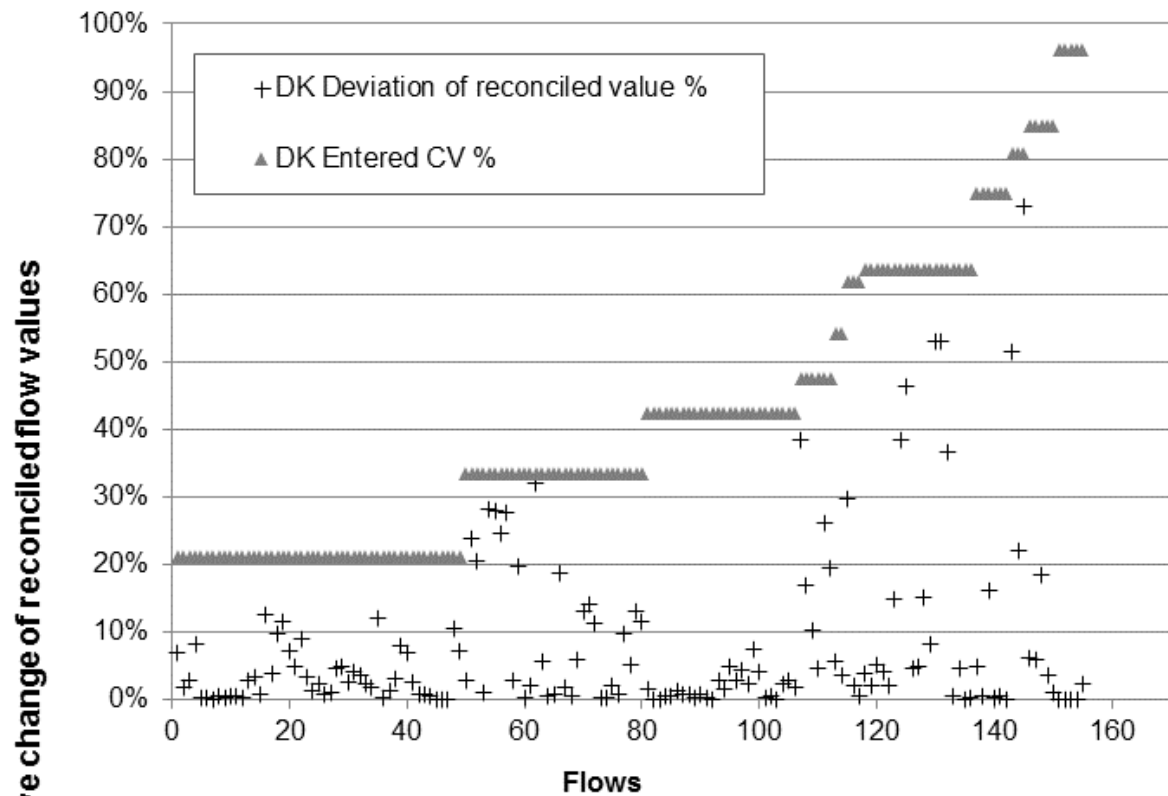


Figure 6b

Figure 6. a) Relative uncertainty (coefficient of variation, or CV) for Denmark (DK) and Austria (AT), as entered and after data reconciliation; b) CV as entered, and relative deviation of reconciled flow values from entered values (in % of the entered value). Each point on the x-axis represents a flow, with flows numbered and sorted from the lowest to the highest uncertainty as entered. Flows calculated in STAN (i.e. no flow values entered) are not shown here.

3.2.3 Aggregation effects

In the case of multi-level studies, such as the Danish multi-regional balance, the aggregation or disaggregation of flows may be of particular importance for the degree of data reconciliation and the resulting flow uncertainties. In the Danish study, in- and outflows to and from regional subsystems are aggregated to flows on the national level, which are also known. However, because the flows on the national level are generally based on the same data sources as regional in- and outflows, treating them as independent estimates may not be fully justified. Therefore, the overdetermination related to the flows on the national level, which are aggregates of regional flows (see Table S.2), causes the resulting uncertainties (after reconciliation) to be lower than in the case of treating the national flows as unknowns (see Section 2.3). In the Danish study, the available information was used to enter values for regional flows as well as aggregate flows in the model, though in many cases the original data stemmed from the same sources. The assumption of independent variables in the reconciliation step may have resulted in lower uncertainties than justified by the original level of information. An alternative approach would be to specify the national flows (with identical data sources as the regional flows) as

unknowns. This would reduce the degree of overdetermination and also the potential for conflicting data. Furthermore, it would result in slightly higher uncertainties of the respective flows.

4. Conclusions

In this study, we showed how choices made in establishing country-scale material balances, i.e. model properties not resulting directly from the physical properties of the systems studied, can influence results. While this may not impact the overall conclusions drawn by such studies, such variance is of particular importance when the values published in an MFA are used in other works. In other words, metadata matter. Regarding the qualitative analysis of system layouts, we showed that the choice of data material is closely linked to the layout of the model and vice versa. We identified the following, more widely applicable, quantitative implications: *first*, conflicting flow data are practically inevitable in country-scale MFA or similar material balances of a complexity comparable to the studies examined in this article. The way in which these conflicts are handled—for example, through the choice of approach for uncertainty characterisation, propagation, and reconciliation—requires transparency. This process reflects the authors’ belief in the data underlying the MFA. The analysis of data reconciliation is an important step to judge the consistency and quality of an MFA model, and to assess both the data and the system. The average extent to which reconciled data deviate from input data in an MFA model provides a useful measure of the data consistency and the quality of the mass balances. The higher the inconsistency between the data given the mass balance constraints, the more data reconciliation becomes analogous to fitting the proverbial square peg in a round hole. A *second* implication of uncertainty assessment being subjective is that uncertainties can be evaluated relative to the uncertainties of other flows in the same model; however, they may not be comparable to the uncertainties specified for a similar flow in another model due to different assumptions and choices

made in each study. *Third*, an iterative approach to quantifying the uncertainty of flow values, i.e. adjusting uncertainty ranges to more closely correlate to the extent of data reconciliation, can ensure the representation of data uncertainty in compliance with the level of information available, while also avoiding overly optimistic or pessimistic estimates of data quality. *Finally*, and linked to uncertainty assessment, the treatment of multi-level flows as independent or not, has an effect on the resulting uncertainty of the flows. This is of particular interest if regional flows are aggregated to national flows, with data for flows on both levels originating from the same sources. As a concluding remark, we stress the need for a transparent system definition and data characterisation in MFA. To better comprehend the results of MFA studies, the authors have the responsibility to document their choices, especially those that are not immediately evident from the model results. This would enable consistent comparisons between key indices (such as recovery rates, etc.). In this article, we have put forward a quasi-standard approach to transparently and systematically document and analyse the choices underlying an MFA model.

Acknowledgements

This research was funded by the Danish Council for Strategic Research (Innovation Fund) as part of the Integrated Resource Management & Recovery (IRMAR) project at the Technical University of Denmark. We would like to thank Oliver Cencic (Vienna University of Technology) for his valuable inputs.

References

- Baccini, Peter, and Paul H. Brunner. 2012. *Metabolism of the Anthroposphere*. 2nd ed. Cambridge (MA): MIT Press.
- Cencic, Oliver, and Helmut Rechberger. 2008. "Material Flow Analysis with Software STAN." *Journal of Environmental Engineering & Management* 18 (1): 3–7.
http://ser.cienve.org.tw/download/18-1/jeeam18-1_3-7.pdf.
- Chen, Wei-Qiang, and T E Graedel. 2012. "Anthropogenic Cycles of the Elements: A Critical Review." *Environmental Science & Technology* 46 (16) (August 21): 8574–86.
doi:10.1021/es3010333. <http://www.ncbi.nlm.nih.gov/pubmed/22803614>.
- Chowdhury, Rubel Biswas, Graham A. Moore, Anthony J. Weatherley, and Meenakshi Arora. 2014. "A Review of Recent Substance Flow Analyses of Phosphorus to Identify Priority Management Areas at Different Geographical Scales." *Resources, Conservation and Recycling* 83 (February 2014): 213–228. doi:10.1016/j.resconrec.2013.10.014.
<http://linkinghub.elsevier.com/retrieve/pii/S0921344913002218>.
- Egle, Lukas, Ottavia Zoboli, Simon Thaler, Helmut Rechberger, and Matthias Zessner. 2014. "The Austrian P Budget as a Basis for Resource Optimization." *Resources, Conservation and Recycling* 83 (February): 152–162. doi:10.1016/j.resconrec.2013.09.009.
<http://linkinghub.elsevier.com/retrieve/pii/S0921344913001900>.
- Fischer-Kowalski, Marina, Fridolin Krausmann, Stefan Giljum, Stephan Lutter, Andreas Mayer, Stefan Bringezu, Yuichi Moriguchi, Helmut Schütz, Heinz Schandl, and Helga Weisz. 2011. "Methodology and Indicators of Economy-Wide Material Flow Accounting." *Journal of Industrial Ecology* 15 (6) (December 31): 855–876. doi:10.1111/j.1530-9290.2011.00366.x.
<http://doi.wiley.com/10.1111/j.1530-9290.2011.00366.x>.
- Hansen, Erik, and Carsten Lassen. 2003. "Experience with the Use of Substance Flow Analysis in Denmark." *Journal of Industrial Ecology* 6 (3): 201–219.
- Hedbrant, Johan, and Louise Sörme. 2001. "Data Vagueness and Uncertainties in Urban Heavy-Metal Data Collection." *Water, Air and Soil Pollution: Focus* 1 (3-4): 43–53.
- Klinglmair, Manfred, Camilla Lemming, Lars Stoumann Jensen, Helmut Rechberger, Charlotte Scheutz, and Thomas Fruergaard Astrup. 2015. "Phosphorus in Denmark : National and Regional Anthropogenic Flows." *Resources, Conservation and Recycling* 105 B (Special Issue: Losses and Efficiencies in Phosphorus Management): 311–324. doi:10.1016/j.resconrec.2015.09.019.
- Laner, David, Julia Feketitsch, Helmut Rechberger, and Johann Fellner. 2015a. "A Novel Approach to Characterize Data Uncertainty in MFA and Its Application to Plastic Flows in Austria." *Journal of Industrial Ecology*. doi:10.1111/jiec.12326.
- Laner, David, Helmut Rechberger, and Thomas Fruergaard Astrup. 2014. "Systematic Evaluation of Uncertainty in Material Flow Analysis." *Journal of Industrial Ecology* 18 (6): 859–870.
doi:10.1111/jiec.12143. <http://doi.wiley.com/10.1111/jiec.12143>.
- . 2015b. "Applying Fuzzy and Probabilistic Uncertainty Concepts to the Material Flow Analysis of Palladium in Austria." *Journal of Industrial Ecology*. doi:10.1111/jiec.12235.
<http://doi.wiley.com/10.1111/jiec.12235>.

- 1 Müller, Esther, Lorenz M. Hilty, Rolf Widmer, Mathias Schluep, and Martin Faulstich. 2014.
2 “Modeling Metal Stocks and Flows: A Review of Dynamic Material Flow Analysis Methods.”
3 *Environmental Science and Technology* 48 (4): 2102–2113. doi:10.1021/es403506a.
- 4 Schwab, Oliver, Ottavia Zoboli, and Helmut Rechberger. 2015. “A Data Characterization Framework
5 for Material Flow Analysis.” *Journal of Industrial Ecology* 00 (0). doi:10.1111/jiec.12399.
- 6 Seyhan, Demet. 2009. “Country-Scale Phosphorus Balancing as a Base for Resources Conservation.”
7 *Resources, Conservation and Recycling* 53 (12) (October): 698–709.
8 doi:10.1016/j.resconrec.2009.05.001.
9 <http://linkinghub.elsevier.com/retrieve/pii/S09213449090001037>.
- 10 Statistics Austria. 2015. “Crops.”
11 [http://www.statistik.at/web_en/statistics/Economy/agriculture_and_forestry/farm_structure_cultiv](http://www.statistik.at/web_en/statistics/Economy/agriculture_and_forestry/farm_structure_cultivated_area_yields/crops/index.html)
12 [ated_area_yields/crops/index.html](http://www.statistik.at/web_en/statistics/Economy/agriculture_and_forestry/farm_structure_cultivated_area_yields/crops/index.html).
- 13 Weidema, Bo Pedersen, and Marianne Suhr Wesnæs. 1996. “Data Quality Management for Life Cycle
14 Inventories-an Example of Using Data Quality Indicators.” *Journal of Cleaner Production* 4 (3-4):
15 167–174.
- 16 Zoboli, Ottavia, David Laner, Matthias Zessner, and Helmut Rechberger. 2015a. “Added Values of
17 Time Series in MFA: The Austrian Phosphorus Budget from 1990 to 2011.” *Journal of Industrial*
18 *Ecology*. doi:10.1111/jiec.12381.
- 19 Zoboli, Ottavia, Alberto Viglione, Helmut Rechberger, and Matthias Zessner. 2015b. “Impact of
20 Reduced Anthropogenic Emissions and Century Flood on the Phosphorus Stock , Concentrations
21 and Loads in the Upper Danube.” *Science of the Total Environment* 518-519: 117–129.
22 doi:10.1016/j.scitotenv.2015.02.087.